APPENDIX C
SAMPLE DESIGN CALCULATIONS: DEVELOPMENT OF DECISION MATRIX (TABLE 5-1)

C-1. The decision matrix will be developed using examples for both the random and stratified random sampling formulas. Initially, use the random sampling formula

$$n = \frac{t^2 s^2}{d^2}$$

as previously defined to estimate the number of epilimnetic samples.

a. The variance s^2 is obtained from the variance component analysis performed on the pilot studies or from the routine sampling program. For this example, assume $s^2 = 100$.

b. Assume a desired precision level d to be, e.g., ±5 μg P/ℓ of a mean phosphorus concentration of 20 μg P/ℓ. Also assume a 90-percent probability of being within the desired precision level of the mean. Since the Student's t value for this probability level varies as a function of sample size, an $n = 30$ is used to initialize the sampling formula, so

$$n = \frac{(1.697)^2_{30,0.10}\ (100)}{(5)^2} = \frac{(2.88)(100)}{25} = 11.52$$

c. Round n to the next larger integer, 12; enter the t table at $\alpha = 0.10$ for $n = 12$ and repeat:

$$n = \frac{(1.782)^2_{12,0.10}\ (100)}{(5)^2} = \frac{(3.175)(100)}{25} = 12.7$$

d. Round to 13, reenter the t table and repeat:

$$n = \frac{(1.771)^2_{13,0.10}\ (100)}{(5)^2} = 12.5 \text{ (convergence at } n = 13)$$

e. To be within ±5 μg P/ℓ of the mean epilimnion P concentration 90 percent of the time, therefore, requires 13 randomly collected samples. Similar analyses performed at transect 10 or 12 for turbidity and chlorophyll, to be within a desired precision level 90 percent of the time, indicated 6 and 10 samples be randomly collected, respectively.

f. If one assumes a fixed cost of \$500 per trip and a per sample analytical cost of \$13 for phosphorus, \$3 for turbidity, and \$20 for chlorophyll, the cost per sampling trip is

$$C(n) = C_o + \sum_{1}^{k} C_i n_i$$

or

$$C(n) = 500 + (13)(13) + (3)(6) + (20)(10)$$

$$C(n) = 500 + 169 + 18 + 200$$

$$C(n) = \$887$$

If one samples 12 times per year, the total annual cost of sampling these three variables in the epilimnion is $(12)(887) = \$10,644$.

C-2. The decision matrix can now be expanded to include other alternatives. If, for example, a desired precision of ±10 μg P/ℓ about the mean P concentration 80 percent of the time is acceptable, the required number of phosphorus samples becomes

$$n = \frac{(1.310)^2_{30,0.20}\ (100)}{(10)^2} = \frac{(1.716)(100)}{100} = 1.7 \sim 2$$

$$\vdots$$

$$n = \frac{(1.638)^2_{3,0.20}\ (100)}{(10)^2} = 2.7$$

$n = 3$ random samples

a. Comparable reductions in the precision and probability levels for chlorophyll result in only three randomly collected epilimnetic samples.

b. If the turbidity precision and probability levels are maintained, the revised sampling costs per trip would be

$$C(n) = 500 + (13)(13) + (3)(6) + (20)(3)$$

$$= \$617/\text{sampling trip}$$

or $7,404 total annual cost (12 × $617) for epilimnetic sampling, a net savings of $3,240.

C-3. If we assume a variance s^2 for the metalimnion and hypolimnion of 144 and 169, respectively, a desired precision of ±5 μg P/ℓ about the mean P concentration with a 90-percent probability level, using the random sampling formula, results in 18 and 20 randomly collected samples from the metalimnion and hypolimnion, respectively. Therefore, to achieve these desired results, 51 random samples would have to be taken throughout the water column.

C-4. The stratified random sampling formula permits a weighting of the variance to reflect its relative importance in the system. While the variability may be greatest in the hypolimnion, the hypolimnion may represent only 30 percent of the entire volume.

a. If we assume the cost of sampling each stratum is constant, the stratified random sampling formula is

$$n = \frac{(\Sigma w_i s_i)^2}{d^2/t^2}$$

as previously defined. If we assume the same desired precision and probability levels, i.e. ±5 μg P/ℓ, 90-percent probability, and weight by stratum volume, then

$$w_{\text{epilimnion}} = 0.40$$

$$w_{\text{metalimnion}} = 0.20$$

$$w_{\text{hypolimnion}} = 0.40$$

$$n = \frac{[(0.4)(10) + (0.2)(12) + (0.4)(13)]^2}{(5)^2/(1.697)^2_{30,0.10}}$$

$$= 15.5$$

$$n = \frac{[(0.4)(10) + (0.2)(12) + (0.4)(13)]^2}{(5)^2/(1.746)^2_{16,0.10}}$$

$$n = 16.4$$

$$n = \frac{[(0.4)(10) + (0.2)(12) + (0.4)(13)]^2}{(5)^2/(1.74)^2}$$

$$n = 16.3$$

$$n = 17$$

The total number of samples using a stratified approach, therefore, has been reduced from 51 to 17.

b. The required number of randomly collected samples per stratum is

$$\frac{n_i}{n} = \frac{w_i s_i}{\Sigma(w_i s_i)}$$

$$\frac{n_e}{17} = \frac{(0.4)(10)}{[(0.4)(10) + (0.2)(12) + (0.4)(13)]}$$

$$= 5.9$$

$$\sim 6$$

$$\frac{n_m}{17} = \frac{(0.2)(12)}{[(0.4)(10) + (0.2)(12) + (0.4)(13)]}$$

$$= 3.5$$

$$\sim 4$$

$$\frac{n_h}{17} = \frac{(0.4)(13)}{[(0.4)(10) + (0.2)(12) + (0.4)(13)]}$$

$$= 7.6$$

$$\sim 8$$

$$n = 18$$

c. Similar computations can be made for each of the other water quality variables.

C-5. The cost of the sampling program again can be computed using the cost formula. As previously described, a decision matrix can be developed that reflects various probability and precision levels to satisfy funding constraints.

C-6. The sampling formula can also be used to assess the loss of precision if fixed numbers of samples are collected at each station. It is, perhaps, not realistic to expect a field crew to collect variable numbers of samples for water quality constituents at each station. If it is determined that six samples will be collected from the epilimnion, four samples from the metalimnion, and eight samples from the hypolimnion, the precision for each constituent can be determined by rearranging the sampling formula as

$$d = \frac{ts}{\sqrt{n}} \quad \text{for the random sampling formula}$$

and

$$d = \frac{(\Sigma w_i s_i)t}{\sqrt{n}} \quad \text{for the stratified random sampling formula}$$

a. For chlorophyll, which will be collected in the epilimnion, six samples will result in a precision, at the 90-percent probability level, of

$$d = \frac{(1.943)(3.5)}{\sqrt{6}} = 2.78$$

$$= \pm 40 \text{ percent}$$

b. Six samples have reduced the precision from ±25 percent for chlorophyll to ±40 percent, but this loss, or gain, in precision is now known and can be calculated.

C-7. In general, use of the stratified random sampling formula results in a more cost-effective, efficient sampling design than using the random sampling approach.